

Partitioning Bipartite Graphs: A Modified Louvain

Emily Diana
emily.diana@yale.edu

Follow this and additional works at: <http://elischolar.library.yale.edu/dayofdata>



Part of the [Other Applied Mathematics Commons](#)

Emily Diana, "Partitioning Bipartite Graphs: A Modified Louvain" (September 25, 2014). *Yale Day of Data*. Paper 8.
<http://elischolar.library.yale.edu/dayofdata/2014/Posters/8>

This Event is brought to you for free and open access by EliScholar – A Digital Platform for Scholarly Publishing at Yale. It has been accepted for inclusion in Yale Day of Data by an authorized administrator of EliScholar – A Digital Platform for Scholarly Publishing at Yale. For more information, please contact elischolar@yale.edu.

Partitioning Bipartite Graphs: A Modified Louvain

Emily Diana

Yale University Class of 2015

Mentors: Geoff Sanders, Van Henson, and Manda Winlaw
Lawrence Livermore National Laboratory

How do we find communities in a graph? How does this change if the graph is bipartite? The Louvain method maximizes links within communities and minimizes those between in order to determine an optimal grouping. Yet, because it may fail when bipartite restrictions are introduced, we have adjusted the null model so as to improve performance in these conditions.

The Louvain Algorithm

- Maximizes modularity coefficient, -1 to 1

$$Q = \frac{1}{2m} \sum_{i,j} \left[A_{ij} - \frac{k_i k_j}{2m} \right] \delta(c_i, c_j)$$

m = number of edges

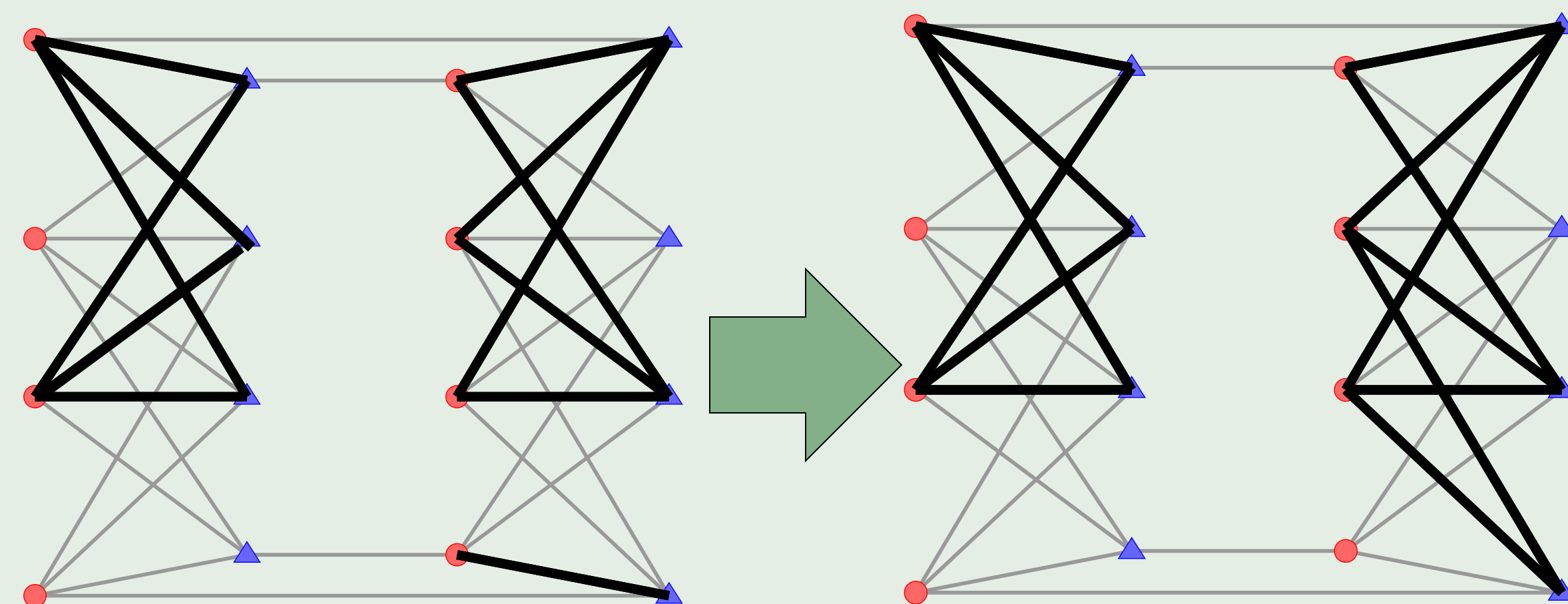
$A_{ij} = 1$ iff nodes i and j are connected

k_i, k_j are the degrees of nodes i and j

$\frac{k_i k_j}{2m}$ indicates probability of edge

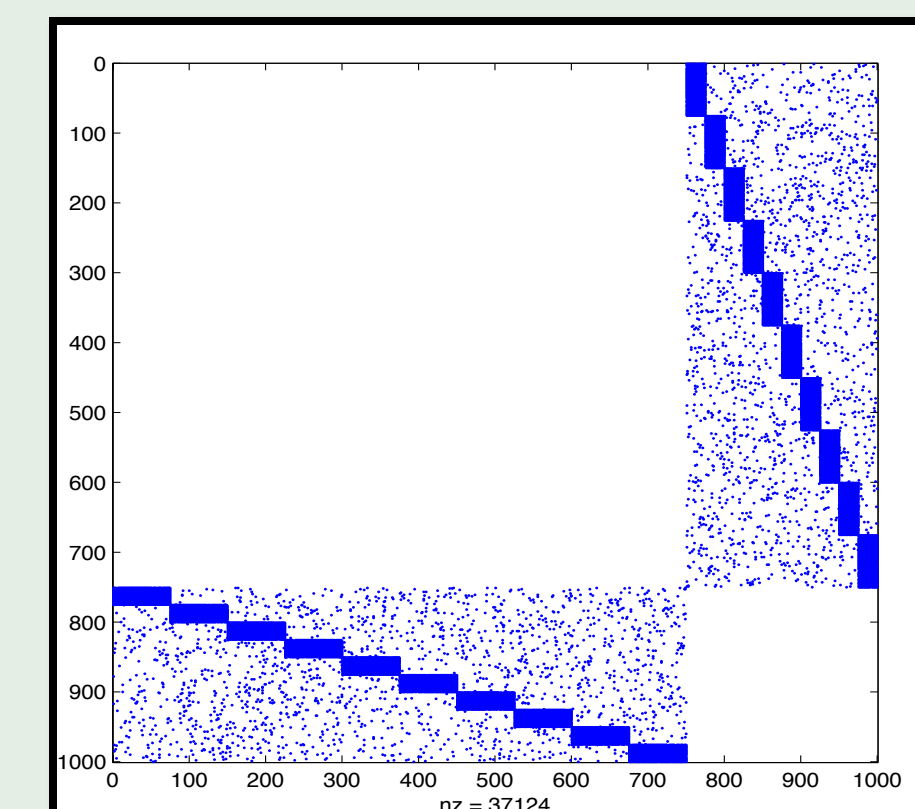
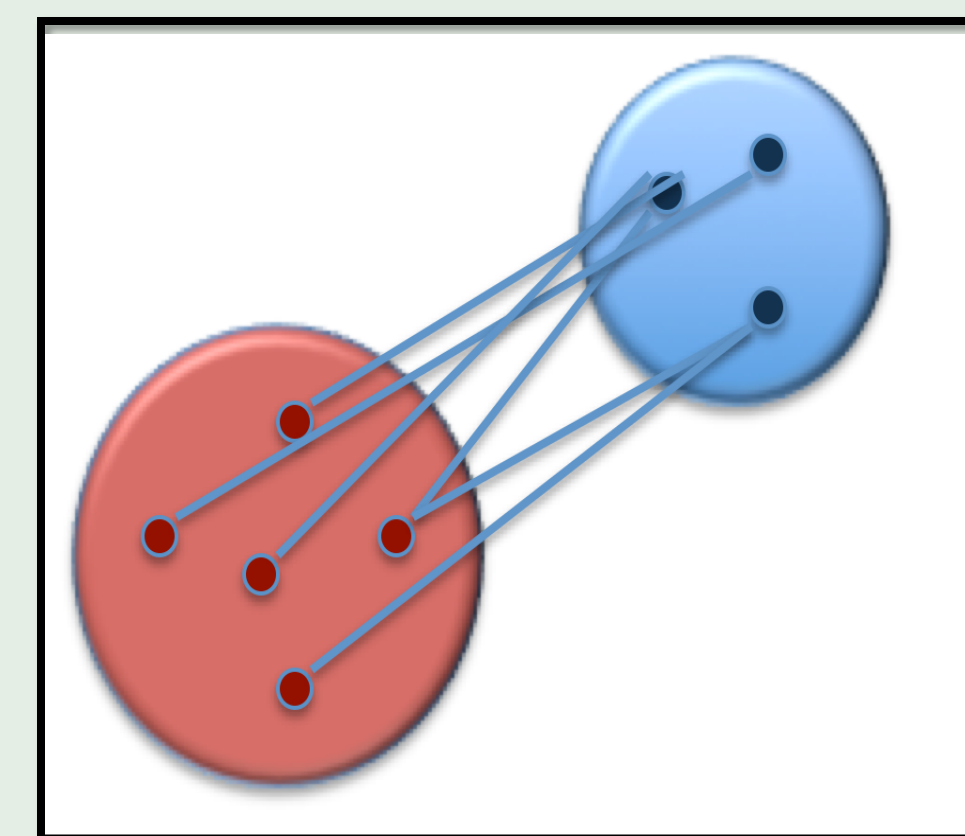
$\delta(c_i, c_j) = 1$ iff nodes i and j are in the same community

Updating Procedure



- Bold edges indicate that the connected nodes are in the same community
- Each node starts in its own community
- What move (or lack of) most improves modularity?
- In next pass, treats each community as a node
- Sweeps through until no moves remain that increase modularity

What if the Graph is Bipartite?



Adjusted Null Model

Original

Bipartite

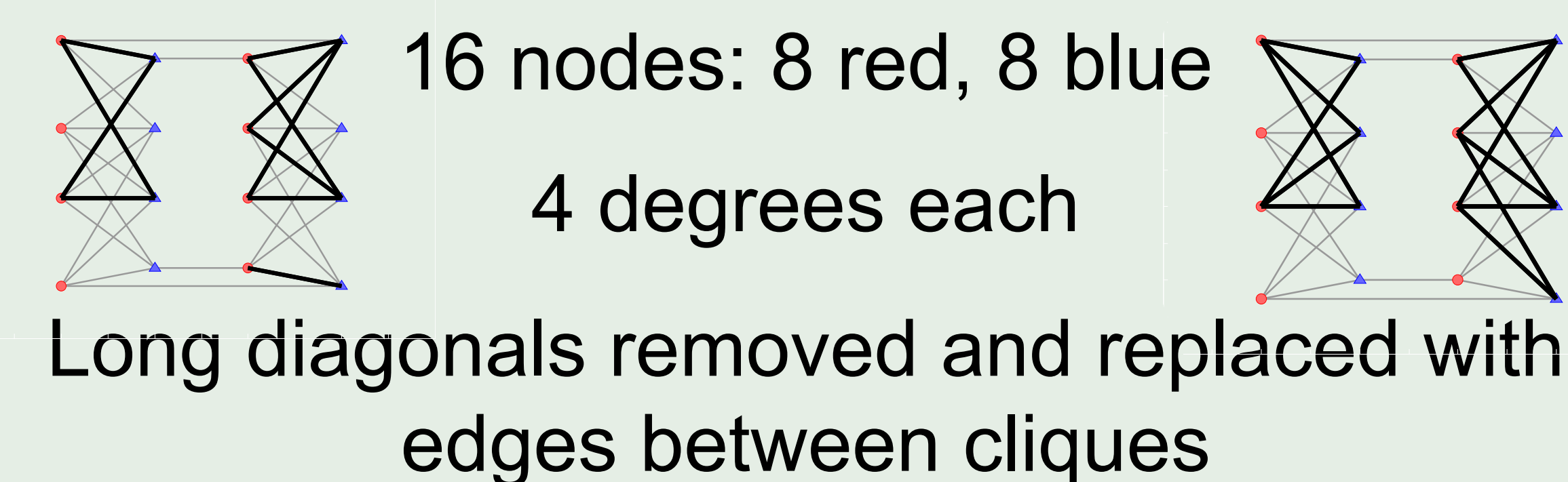
$$A = \begin{bmatrix} \frac{rr^t}{2m} & \frac{rb^t}{2m} \\ \frac{br^t}{2m} & \frac{bb^t}{2m} \end{bmatrix} \quad A = \begin{bmatrix} 0 & \frac{rb^t}{m} \\ \frac{br^t}{m} & 0 \end{bmatrix}$$

- Doesn't consider impossible connections
- Half the number of edges
- Bipartite modularity coefficient now defined as:

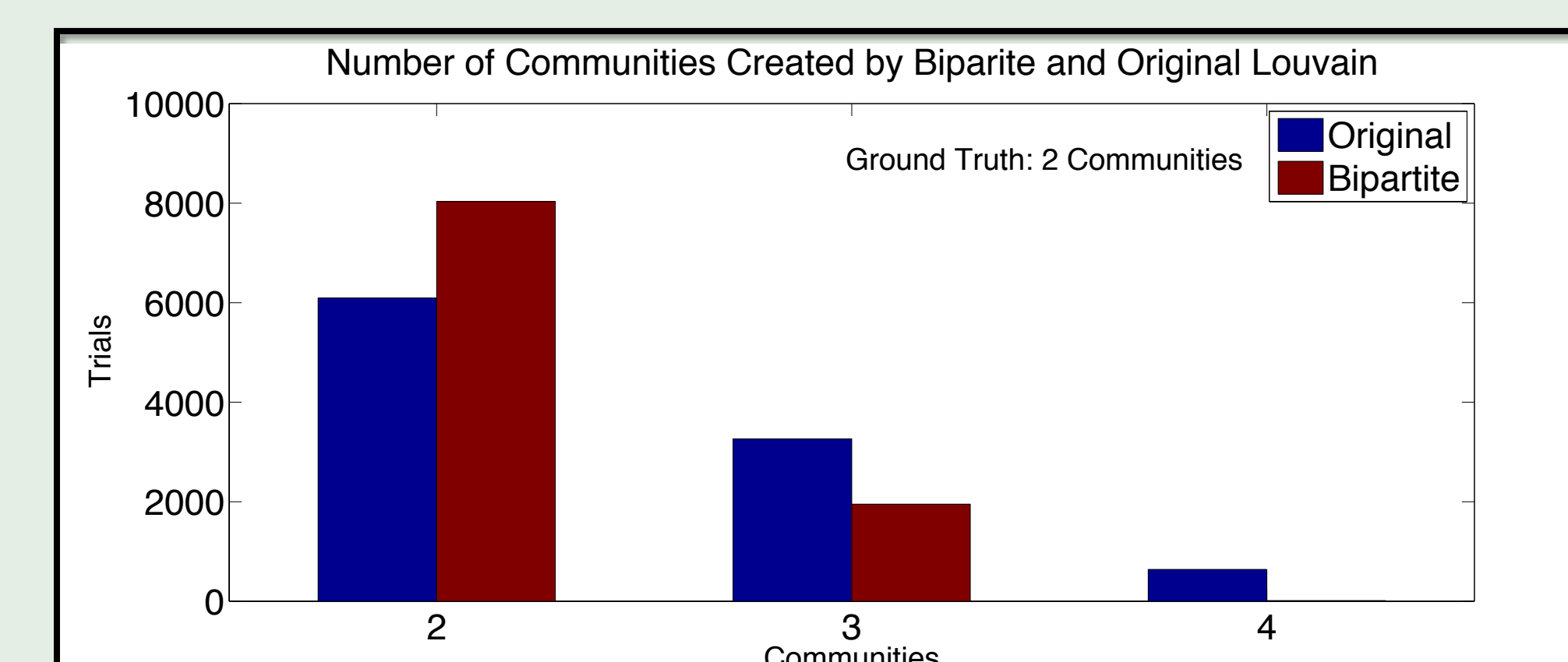
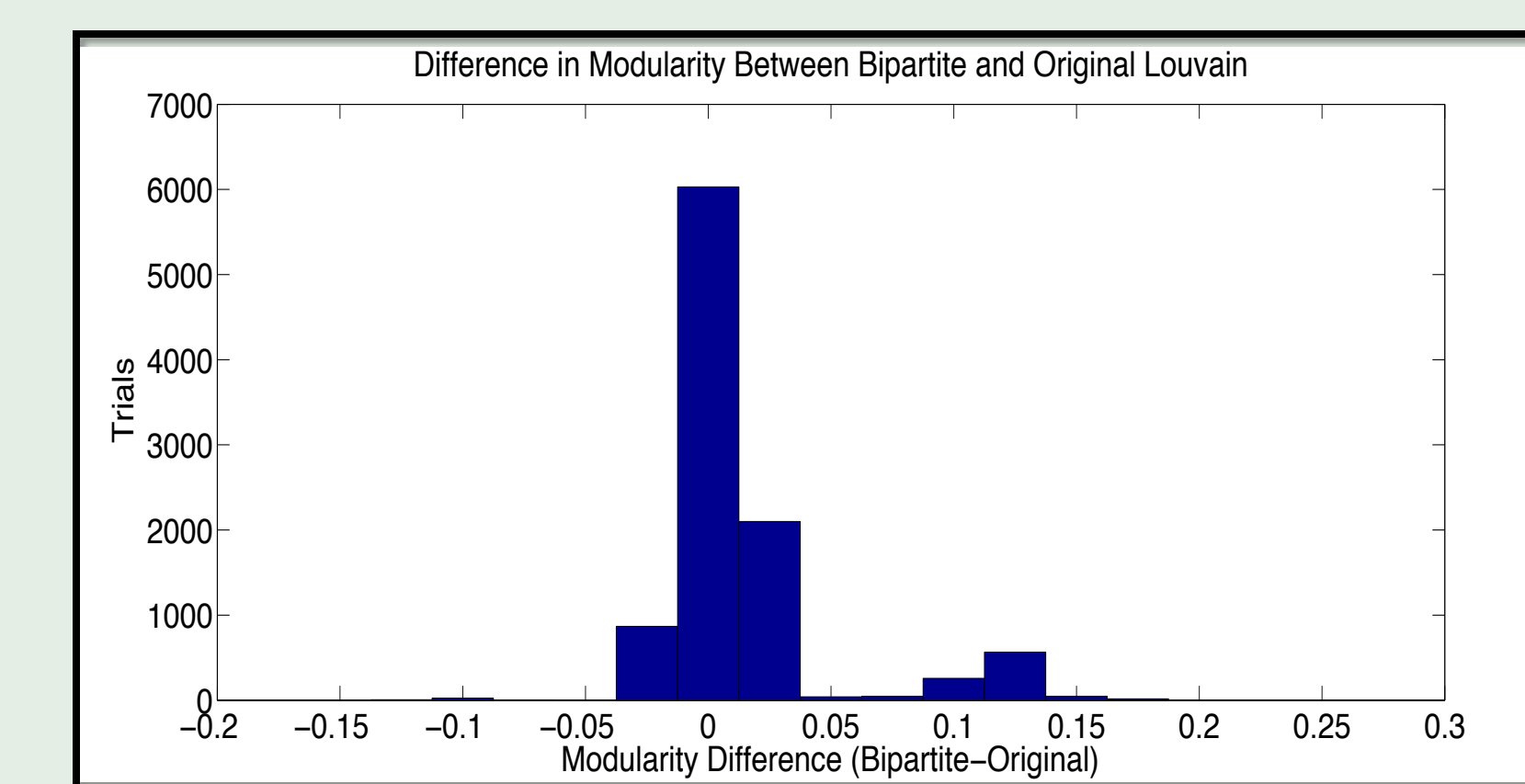
$$Q = \frac{1}{m} \sum_{i,j} \left[A_{ij} - \frac{k_i k_j}{m} \right] \delta(c_i, c_j) \xi(g_i, g_j)$$

$\xi(g_i, g_j) = 1$ iff nodes i and j are of different colors

A Small Example



10000 permutations of the nodes



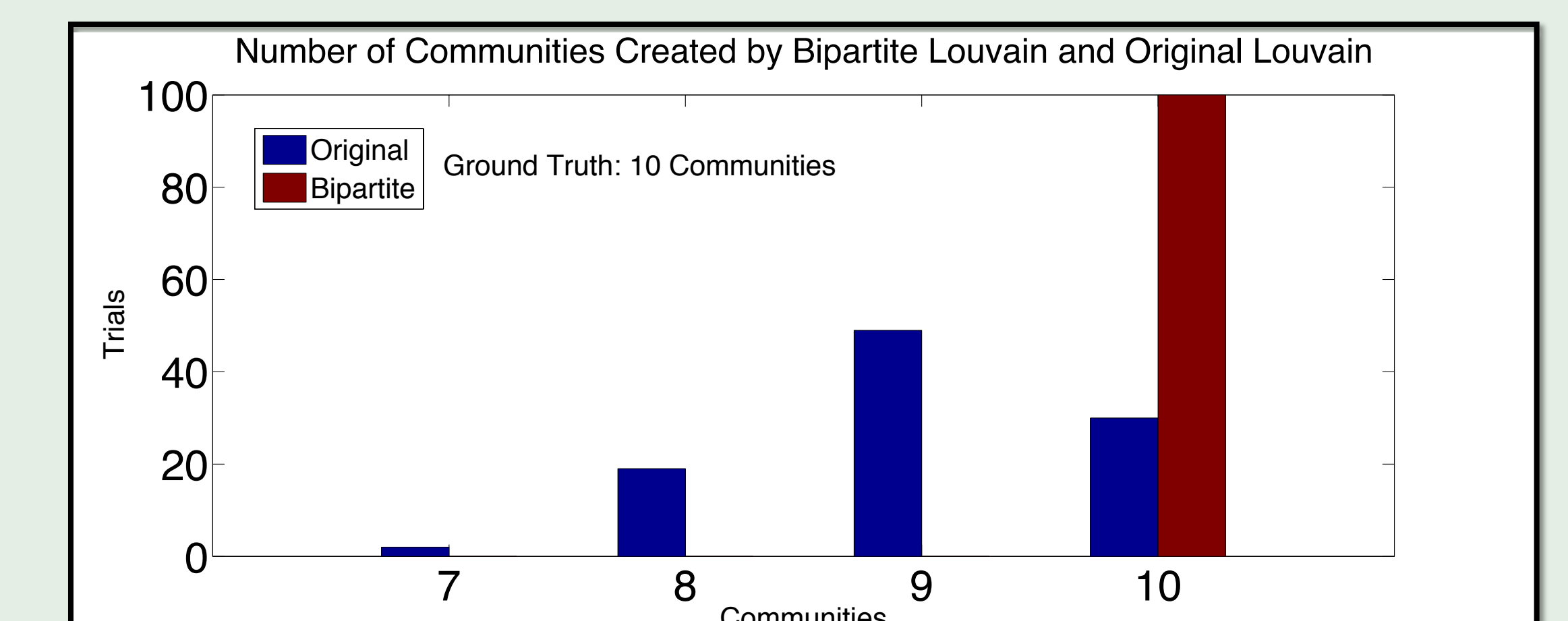
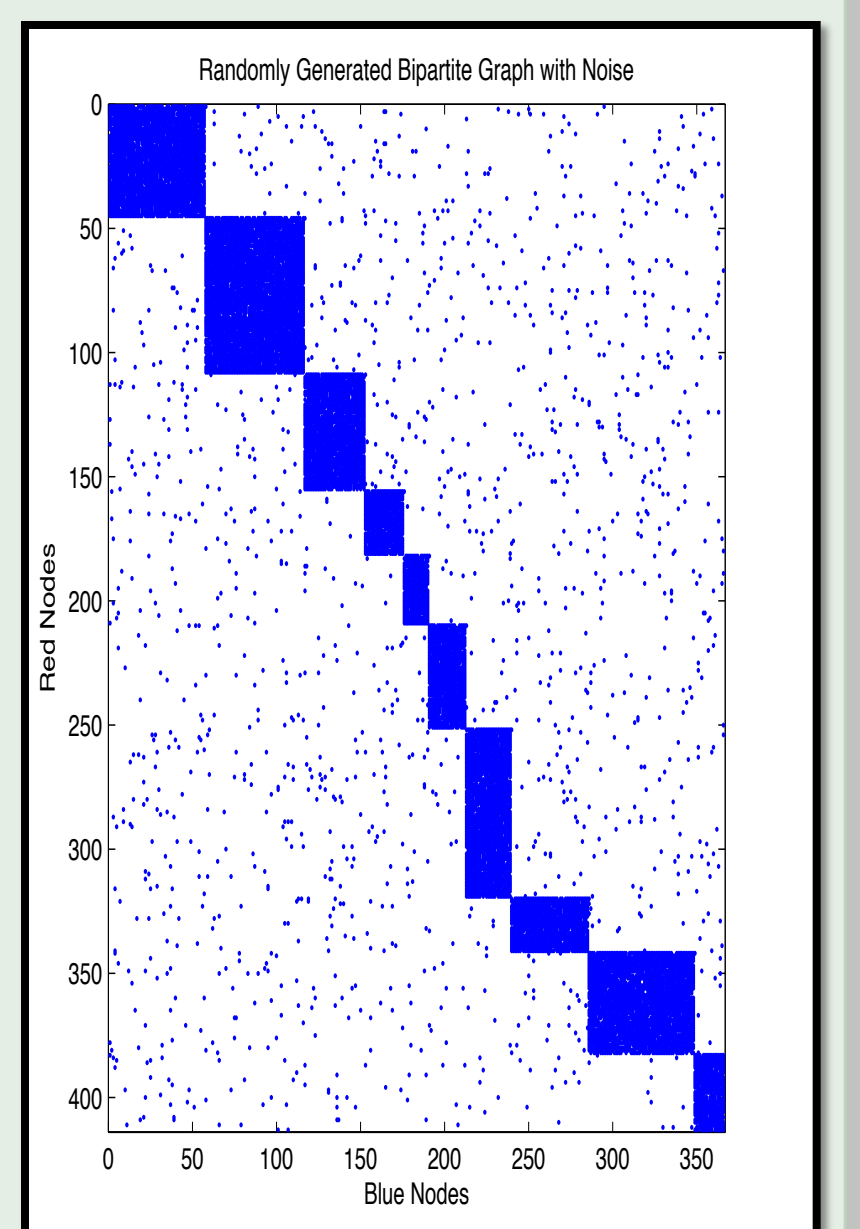
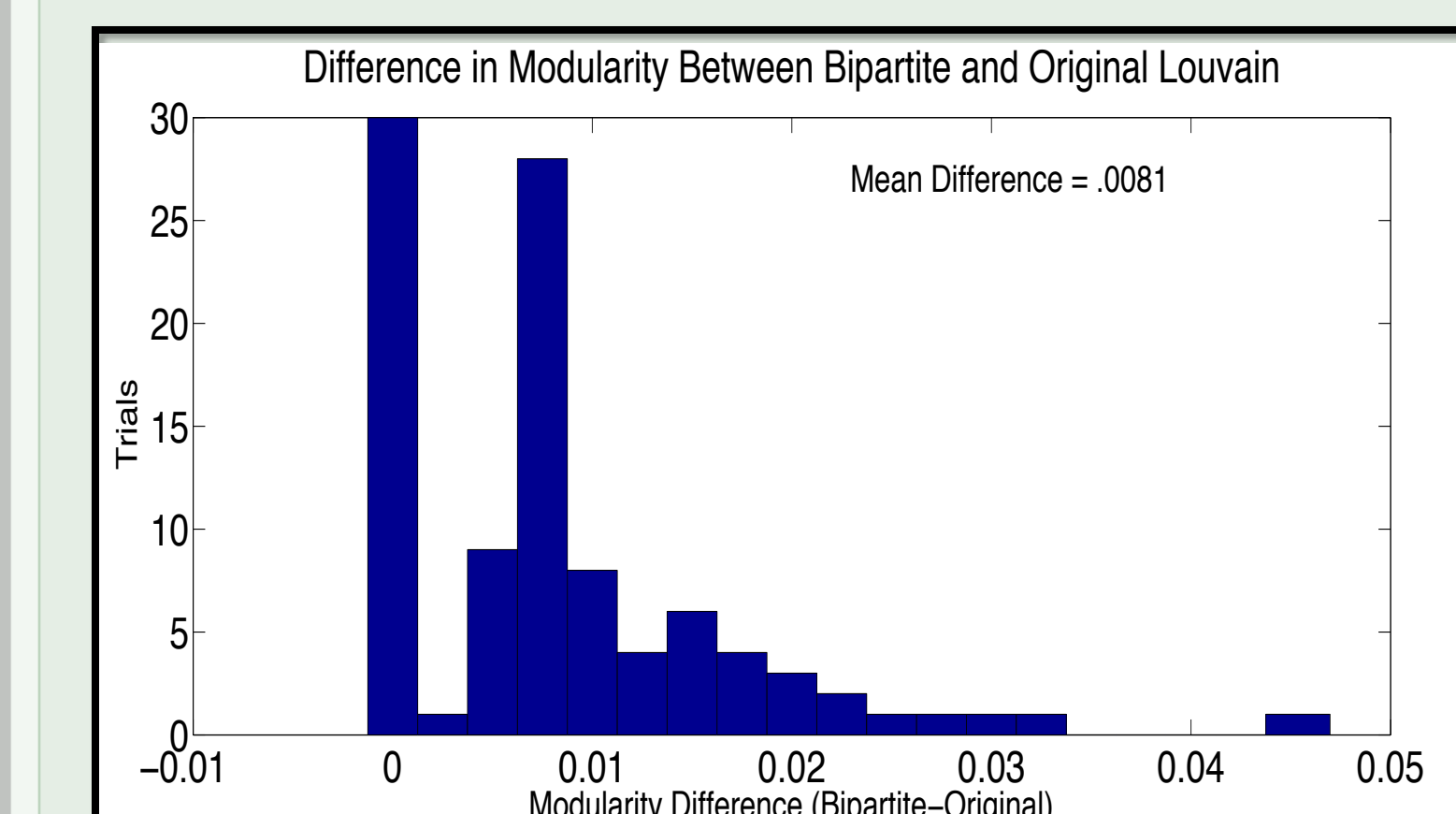
Results: Bipartite Louvain improved modularity by an average of **.0145** and correctly detected two communities **80%** of the time as compared to 61% by standard Louvain.

A Large Example

Randomly generated graph with noise

779 vertices: 413 red, 366 blue

100 permutations of the nodes



Results: Bipartite Louvain improved modularity by an average of **.0081** and correctly detected 10 communities in every case as compared to 30% by standard Louvain.

Conclusion and Future Work

Our Bipartite Louvain is more robust with respect to permutations of vertices than the standard Louvain. For our synthetic examples, Bipartite Louvain typically yields a higher modularity and uncovers the ground truth communities with a higher probability. In the future, we will examine real world data sets with our modified algorithm.

References

- Blondel, Vincent D., Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. "Fast unfolding of communities in large networks." 25 July 2008.
- Larremore, Daniel B., Aaron Clauset, and Abigail Z. Jacobs. "Efficient inferring community structure in bipartite networks." 10 July 2014.
- Barber, Michael J. "Modularity and community detection in bipartite networks." 7 December 2007.